Sensor Failure Prediction and Reliability

[1]M Kowsigan, [2]Sanjay S, [3]Rohit Ravi, [4]Kanagachidambaresan G R
[1]Assistant Professor, Department of CSE, SRM Institute of Science and Technology
[2]B.Tech, Department of CSE, SRM Institute of Science and Technology
[3]B.Tech, Department of CSE, SRM Institute of Science and Technology
[4]Associate Professor, Department of CSE, Vel Tech University
Email: [1]mkowsigan@gmail.com

Article Info
Volume 83
Page Number: 1461 - 1467
Publication Issue:
March - April 2020

Article History
Article Received: *24 July 2019*
Revised: *12 September 2019*
Accepted: *15 February 2020*
Publication: *14 March 2020*

Abstract

Sensors are vital components of the manufacturing industry due to their various applications in the production, chemical and healthcare industries. Therefore, to obtain better results and facilitate monitoring, it is of utmost importance to spot anomalies and faults in sensors in the quickest possible way, and to test their reliability and functionality upon detection and subsequent repair. We use sensor data available from industrial usage, and create an Artificial Neural Network system to carry out prognostics for the next immediate occurrence of a sensor failure, by detecting anomalies in advance in its data. In industry 4.0, Recurrent Neural Network systems utilise the fact that readings of a sensor are not arbitrary, but rather correlate to a regular pattern. Under this category, the Long Short-Term Memory technique is implemented in the process to generate maximum accuracy. General machine learning algorithms like Regression Tree and Support Vector Machine are chosen for comparison with Artificial Neural Networks to establish maximum possible efficiency and accuracy for the latter in the predictive process. These methods have many modes of application in fields of prediction and classification due to which they have been specifically chosen to perform the same tasks that the Artificial Neural Network is exposed to. Therefore, using these predictor and clustering based methods on the available datasets, the system generates a new predictive model that'll help determine the occurrence of a sensor failure beforehand.

Keywords: Industry 4.0; Machine Learning; Prognostics; Artificial Neural Networks; Long Short-Term Memory

1. Introduction

Sensors are used in a massively wide spread network of the industry, ranging from civil and data-based infrastructure, automation, medical machinery, and many other departments of science and security [20, 21]. The healthcare, chemical, production and satellite industries of today especially, require and use sensors on a consistent basis to ensure proper functioning of equipment and the necessary machinery as a whole [6]. This makes it essential for the production of better results and effective monitoring of these sensors and machines in the manufacturing industry. Therefore, the detection of noticeable inconsistencies in a sensor's recorded and collected data is necessary to predict sensor failures in advance. This is done in order to test its functionality and reliability upon detection, leading up to subsequent repair. Due to large scale production accompanied with the introduction of industry 4.0 where countless sensors are in use, it is very time consuming and laborious to ensure their functionality and working. Sensor failure and defects can still be left undetected despite a thorough and periodic sensor maintenance system [5]. Furthermore, to enhance their impact, more efficient methods are sought out by businesses to carry out their operations. With the smart industry's rise and relevance, accurate prediction of equipment or process failure may help improve decision making in regards to performance maintenance, cost minimisation and workload reduction.

Existing techniques that include Regression Trees, Random Forest and Support Vector Machine algorithms [13], are chosen for comparisons with Artificial Neural Networks to establish a clear improvement in efficiency and accuracy of the predictive model. Since these classical machine learning techniques are widely used for predictive analysis in various application areas [5, 10], they are therefore chosen to be compared with the ANN techniques to which the datasets are exposed. In real-life, input and output relationships are very complex, due to their non-linear nature. ANNs provide a solution through their ability of learning the specific nature of those existing non-linear relationships, and thus modeling their complexity. This is really important because, unlike many other prediction techniques, there are no restrictions present in the ANN on the input variables.

The main objectives of this process are to propose an approach that uses the available sensor data from a test device to generate a system that builds a dataset and trains a neural network model that can predict the occurrence of a failure in sensors [14]. It is therefore essential that a time period can be determined for the maintenance of sensors and its peripheral systems.

Thus, these prognostics help determine predictive failure of the sensor system, that in turn helps maintain it. The purpose of the paper is also to establish that the ANNs show a greater accuracy in predictive analysis when compared to the general machine learning methods used in earlier cases. Different classes of ANNs are used to observe and compare the best possible system that can be generated out of the given models.

2. Related Works

In the production industry, predictive failure analysis has been gaining enormous traction off late, and a lot of study has been conducted in this field as well. In the past, data mining techniques were used in the process of predictive analysis. Detailed accounts of parts of the proposed methodologies have also been used in the oil refinery industry. This technique was introduced in 2011 [1] to detect sensor failure by looking at its anomalies, thereby generating a means to predict the next occurrence of a malfunction. To detect these inconsistencies, the authors made use of the concepts from the fuzzy logic algorithms [19]. For evaluating the model, the same system was put to use to attain a promising result. A newly derived time series model had to be generated from the original time series models that were available. The newly generated time series had to be plotted against, and this was done with the help of regressive and autoregressive prediction models. With its help, any detectable sensor failure could be predicted and the models' results were compared. The work of the authors showed that the developed model helps both the operators as well as the engineers of oil refineries.

Sensor maintenance in general is a systematic approach of period checks and renewals of the sensors' working, and also replacing them if found malfunctioning. It was however deduced in a paper in 2016 [4] that optimally preparing for the replacement can decrease the cost incurred on its maintenance. The article aimed to suggest a predictive maintenance system for sensors that were monitored and detected for premature failure based on their recorded readings. Different approaches were introduced, first of which dealt with feature extraction and classification. Then there was an approach that based its model on time series. The final approach used auto encoders to detect anomalies.

In 2016 [16], another paper was detailed by its premise on predicting sensor failure which made use of machine learning techniques. The following was assumed, - a light sensing device with id 1053 emitting different types of messages where the message type indicated a specific condition (low-ambient-light / high-ambient-light / twilight etc.) of the environment. This research was carried out in the domain of a smart building and smart city where we have millions of sensors running round the clock.

Further strides were made in this particular field of interest, which was further studied in 2018 [15, 18, 19] when ambient systems were developed to help assist senior citizens in staying active to help them stay independent during their old age. This was done by observing their behavior and enabling the necessary means to capture digitally any signs of health issues prior to the actual event. The non-intrusive sensors however generated negative results leading to an ambiguous interpretation as to the status of elderly people. Their research presented an analysis on the sensor fault detection and its tolerance.

Optical networks utilise optical fibre cables as the major means for communication and conversion of data between two nodes. A paper regarding this was published in 2017 [3], which proposed a system for failure prediction in optical networks. Machine learning techniques like the Support Vector Machine were used, along with Double Exponential Smoothing. The proposed plans on models in optical networks were studied to forecast the risk of failure of the equipment.

In the most recent development on the usage of primary machine learning algorithms affecting the accuracy levels of failure prediction, machine learning's predictive abilities were explored in 2019 [17, 22]. To improve the accuracy of predicting failure, a number of machine learning algorithms were included like those of Support Vector Machines, Classification and Regression Trees, and Random Forest, in addition to Linear Discriminant Analysis and also k-nearest means. Using time series and machine learning, they managed to create a prediction model that performed comparison-based tests on its accuracy. Experimental results suggest that when using SVM, their model's prediction accuracy was more

March-April 2020
ISSN: 0193-4120 Page No. 1461 - 1467

accurate and efficient in relation to other algorithms when predicting failure.

3. Proposed Work

In the proposed system, we have introduced the concept of Artificial Neural Networks (ANN). The derivative of an ANN model, the Long Short-Term Memory technique is brought into the process. This falls under a Recurrent Neural Network based model, which is used to ensure that the readings obtained from a sensor to form datasets are not just randomly generated, but instead infer regular cycles of reference. This proposed system seeks to establish the accuracy, efficiency and performance level generated with the Artificial Neural Network model that has been trained, in comparison with other general machine learning techniques used in prior works to carry out prediction and estimation processes. Artificial Neural Networks based models can be used in various fields of applications, along with the already existing Support Vector Machine, Random Forest and Regression Tree algorithms.

Due to most models only being able to learn and gather information from short-term models, a different approach is needed for the long-term models. This is where the ability of an LSTM network is put to use as it learns from such dependencies that are not just short-term based. In recurrent neural network models, chain like structures is formed. Here, there exists a repetition of neural network based modules upon which the RNN takes shape. On the other hand, standard RNNs will have a very rudimentary structure of these recurrent modules. But in an LSTM, the repeating module of its structure is very different despite it also exhibiting a chain like property, like that of a single tanh layer. The difference lies in the fact there are four different neural network layers coexisting in a specific way, compared to just the single neural network layer being used.

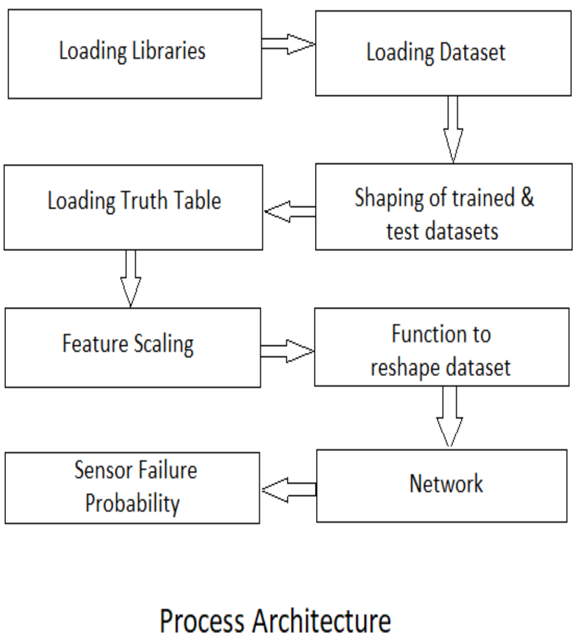

Figure 1: Architecture Diagram

Usage of Artificial Neural Networks [9] eradicates various restrictions and limitations that are present in other machine learning techniques. As discussed, input and output relationships are very complex due to their non-linear nature. ANNs help provide a solution as they can learn nature of such existing non-linear relationships, by modeling their complexity. The advantage that this model provides lies in its non-imposition of any rules on the input obtained, unlike the other numerous prediction techniques. The ANN model does this by first learning about the initial inputs, and also from them. They read into the relationships existing in these initial observations, and help correlate complex and hidden relationships with the hidden data. This in turn helps generalize the model which facilitates prediction based on the hidden data as well.

ANNs also tackle hetero-skedasticityi.e. highly volatile data with a non-constant variance. It achieves this by learning from discrete relationships in the recorded data, and simultaneously not having the need for having to impose any fixed or pre-determined relationships in the data. In places where data volatility is very high, like those of stock markets, time series based financial forecasting can be done. Knowing that real system measurements would need to be accounted for, the influence of variables like pressure and temperature, and other internal and external factors could facilitate further research in the construction of better engineered fault prediction systems as the behavior of these variables would also be different. Smart industries benefit from the further advancements and implementation of these studies [10].

Due to the continuous nature of the fields with which it is to be computed, its applications include language modelling, machine translation, speech and handwriting recognition, image capturing, and so much more.

4. Implementation

Neural networks are used with the help of biological neurons that help create and approximate an unknown function based on its dependency on several inputs. Single and multilayered neural networks show that it is possible to compute a continuous output as opposed to only a step function as formerly believed. Single layered neural networks have only a singular link between input and output, and hence contain only two such layers. Multilayer neural networks act as a connection of multiple layers of inputs and outputs that are already layered. In an ANN, the function of neurons in one layer is a direct coefficient of the actions of the same network of neurons located in another layer.

In figure 2, a Recurrent Neural Network's neuron structure is shown.

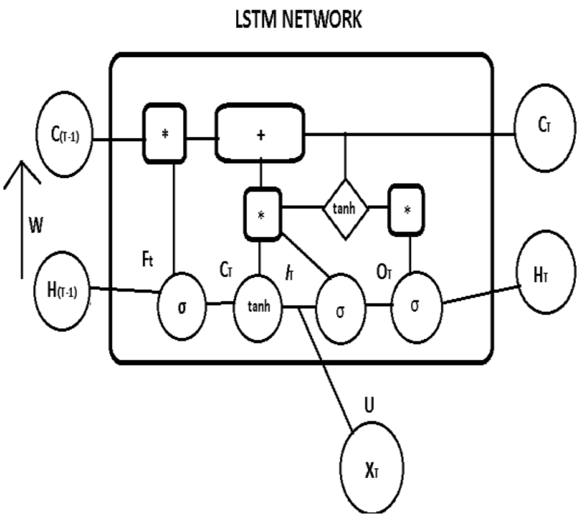

Figure 2: LSTM Network

From the diagram, we infer -

Xt –Input Vector,
Ht-1 – Previous Cell Output,
Ct-1 – Previous Cell Memory,
Ht – Current Cell Output,
Ct- Current Cell Memory,
Ct= Ct + (It*C`t),
Ht = Tanh(Ct)

Each of the neuron's connections has a weight associated with it that represents its electrical signal's intensity that in turn is proportional to the relevance of the available information.

LSTMs consist of several such architectural units. Its general architecture contains a memory unit, called the 'cell', and its three gates, otherwise known as regulators. These gates are responsible for the data flow inside an LSTM's structure. This structure holds an input, an output and a forget gate. Several types of these units exist where there's an absence of either one of more of these gates, or sometimes just the presence of all of them. A Gated Recurrent Unitisa good example of this due to the lack of an output gate in it.

In its input, the memory block, i.e. the cell maintains the nature of the elements' dependencies amongst themselves. The degree of information that is allowed to get to the cell is controlled by the input gate. The forget gate meanwhile is used to enforce the degree to which a value stays in the cell. Lastly, the output gate is responsible for the control of the value in the cell which is used to compute the output activation. The logistic sigmoid function is used as a means to have a general activation function of an LSTM network.

The connections that are present in and out of LSTM gates consist of some recursive functions. During the training of the datasets, the weights of the functions are analysed, which in turn helps in the operation of the existing gates. This induces the use of optimisational algorithms. Gradient descent is one such algorithm used on the sequences of the training datasets. In addition, back propagation is another integral part of the optimisational necessity within the LSTM model. When back propagation of the error values takes place from the output layer along with simultaneously making sure the error stays within the cell inside the LSTM, this creates a feedback loop of errors that constantly feeds itself back into each of the gates of the LSTM unit, until the value is terminated.

4.1 Description of the dataset

The dataset that we have used in the project has been acquired from Kaggle. This dataset was schemed at the International Conference on Prognostics (PH08).

Index	Data fields	Type	Descriptions
1	id	Integer	aircraft engine identifier, range [1, 100]
2	cycle	Integer	time, in cycles
3	setting1	Double	operational setting 1
4	setting2	Double	operational setting 2
5	setting3	Double	operational setting 3
6	s1	Double	sensor measurement 1
7	s2	Double	sensor measurement 2
...	...		
26	s21	Double	sensor measurement 21

Figure 3: Dataset with various parameters

The dataset consists of various parameters involved in the sensors present in aircraft machine. We then segregate the dataset into training and testing dataset. A quintessential part of evaluating data mining models is in the isolation of data into train dataset and test datasets. By dividing a dataset into training and testing sets, the bigger portion of the dataset is used for training, while a comparatively smaller portion of the data is used for testing. The training dataset contains time series with "cycle" as the time unit. Every time series can be presumed as being derived from a distinct machine of the similar sort. Every machine is presumed to begin with distinct degrees of primary wear and manufacturing variation. In this dataset, the machine is expected to function properly at the beginning of each time series cycle. During operating cycles, it starts to degrade. The degradation continues and increases significantly. As the threshold is reached, then the machine is terminated for upcoming operation. The last cycle in each time series is the failure point of the respective machine.

March-April 2020
ISSN: 0193-4120 Page No. 1461 - 1467

4.2 Building an ANN model

Before building the ANN model, Feature Scaling is performed on the training and testing dataset. Feature Scaling is a mark of Data Pre- Processing that is enforced to features of data. It helps to normalize the data. It helps in boosting up the algorithm's evaluation.

The dataset has to be reshaped to be able to generate sequences and labels. This is performed by creating an LSTM Neural Network. We then train and test the dataset in the LSTM Network.

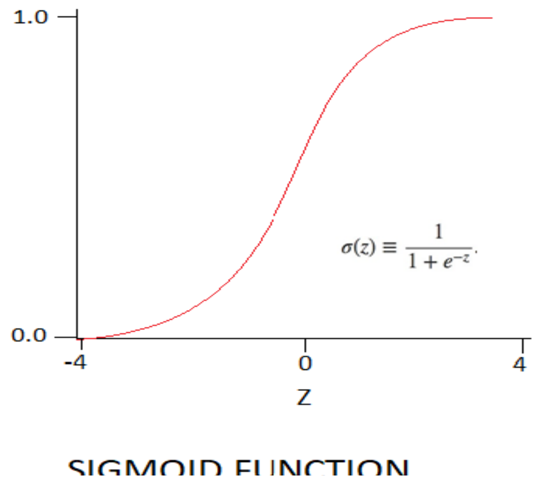

Figure 4: Sigmoid Function

The parameters involved in the dataset are in different formats. A Neural Network can only perform linear mappings from inputs a to the outputs b. Thus, we use an activation function. The use of activation function is to add non-linear property to the function, which is a neural network. So, we will use Sigmoid activation function. The result of the activation function is basically the predicted output for the input features.

4.3 Model Accuracy

Now we can test the model with the test data. The model accuracy on the test set is generated, and it is then compared to the training accuracy. According to the definition, the training accuracy should be idealistic since the model was optimized for those particular observations. The test set accuracy is more conventional, and simulates how the model was intended to be used to predict forward in time. This is the value we should use for describing how the model functions.

5. Results

The models using the Long Short-Term Memory Neural Network were trained and tested. The results that were accomplished during this process had the maximum accuracy obtained when compared to other machine learning algorithms. In the present system, Logistic Regression is used in order to calculate the accuracy and it turned out to be 90.71%. The accuracy of Neural Network model obtained on the test data is 98.56%. The dataset was trained in 19504 samples and validated on 1027 samples. It obtained a validation accuracy of 80.96%. Since the validation accuracy is far more lesser than training accuracy (Figure 5), the model is Overfitting, which explains the fact that the model performs very well and has better prediction results.

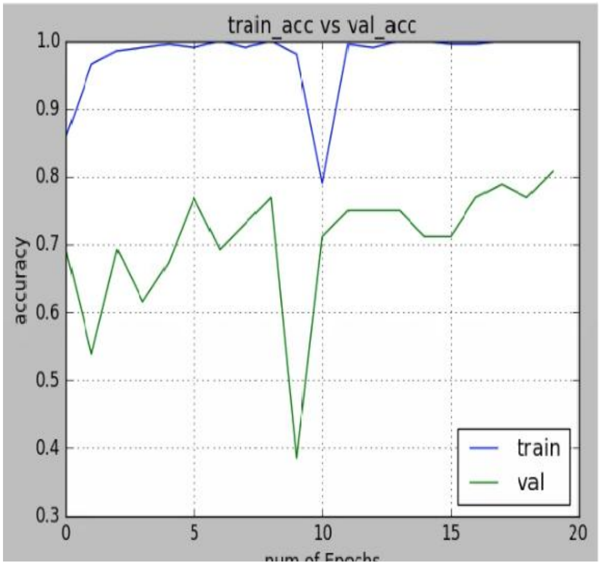

Figure 5: Accuracy vs. Epochs Graph

In Figure 6, the validation loss is higher than the training loss, thus the model is overfitting.

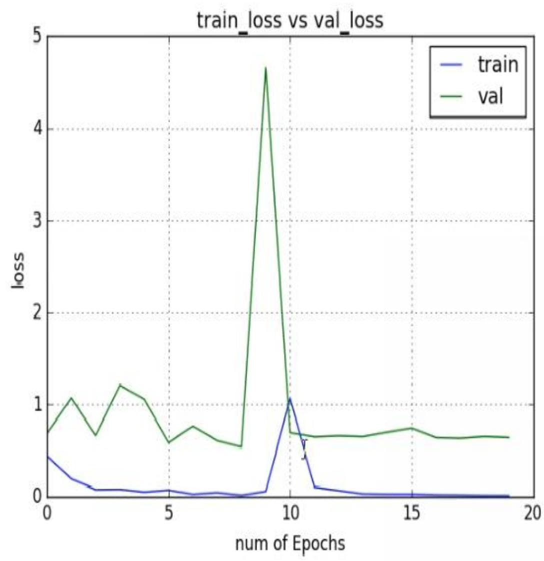

Figure 6: val_loss vs. train_loss

As established, overfitting or underfitting the data explains the cost of poor performance analysed in machine learning. Overfitting is a model's understanding of the detail and noise existing in the training data which might impact its performance of the model negatively, whereas underfitting doesn't model the training data at

March-April 2020
ISSN: 0193-4120 Page No. 1461 - 1467

all, nor does the new data get generalised.
A confusion matrix denotes the performance of the model on the test dataset.

```
Accuracy of model on test data:  0.9856879039704525
Confusion Matrix:
 [[12637    27]
 [  159   173]]
```

We also predicted the Failure Probability of the machine in 30 days and it resulted in an estimated value of 0.03939.

6. Conclusion

The study carried out in this paper presents a practical system towards the prediction of failure in sensors using machine learning, and Artificial Neural Networks based on the 'Long Short-Term Memory' approach. Anomalies are observed, and thereby faults are detected in advance using the results obtained upon the use of the required tools on the sensor data available. The proposed method usually outperforms other machine learning models as it deals with non-linearities automatically. Performance level and accuracy percentage are better when the recurrent ANNs are used. It is discovered that the standard machine learning techniques based on classification and regression are used only for short term predictive analysis. The general analysis in predicting short term proves Artificial Neural Networks, Regression Trees and Random Forest techniques to be equal. However, when it comes to long term predictions, the results obtained from the primary machine learning tools were inaccurate. This is where Artificial Neural Network provided superior results. ANN models were able to generate a higher accuracy percentage for prediction in both the average and longer cycles.

The proposed system looks to provide a source and reference for future studies and research. Research on predictive analysis necessitates technological development that helps provide maximum accuracy using the prediction model and guarantees better results, which in turn satisfies the pursuit of creating time efficient and high performance-based equipment in the industry. Thus, a novel approach or model is introduced to predict sensor failure and ensure their reliability.

References

[1] Mahmoud Reza Saybani, Ying Wah, Amineh Amini and SaeedReza Aghabozorgi Sahaf Yazdi."Anomaly detection and prediction of sensors faults in a refinery using data miningtechniques and fuzzy logic", November2011.

[2] Nancy E. ElHady and Julien Provost. "A Systematic Survey on Sensor Failure Detection and Fault-Tolerance in Ambient Assisted Living",2018

[3] Gongora, W.S.; Silva, H.V.D., Goedtel, A., Godoy, W.F., da Silva, S.A.O, "Neural approach for bearing fault detection in three phase induction motors", August2013.

[4] F. Dikbiyik, M. Tornatore, and B. Mukherjee, "Minimizing the Risk From Disaster Failures in Optical Backbone Networks," J. Lightwave Technol, 2014.

[5] Plante, T.; Nejadpak, A.; Yang, C.X., "Faults detection and failures prediction using vibration analysis", Novermber2015.

[6] Klimberg, R. K., McCullough, B. D., SAS Institute, "Fundamentals of Predictive Analytics with JMP", 2016.

[7] Marco Cerliani. "Predictive Maintenance: detect Faults from Sensors with CNN", March2019.

[8] Xiaoyi Sun, Krishnendu Chakrabarty, Ruirui Huang, Yiquan Chen, Bing Zhao, Hai Cao, Yinhe Han, Xiaoyao Liang, Li Jiang, "System-level hardware failure prediction using deep learning", June 2019.

[9] Shang-Yi Chuang, Nilima Sahoo, Hung-Wei Lin, Yeong-Hwa Chang, "Predictive Maintenance with Sensor Data Analytics on a Raspberry Pi-Based Experimental Platform", September 2019.

[10] Gustavo Scalabrini Sampaio, Arnaldo Rabellode Aguiar Vallim Filho, Leilton Santos da Silva and Leandro Augusto da Silva, "Prediction of Motor Failure Time Using An Artificial Neural Network", October 2019.

[11] Haykin, "S.O. Neural Networks and Learning Machines", 3rded;NJ, USA,2008.

[12] Brand, L., Patel, A., Singh, I., Brand, C., "Real Time Mortality Risk Prediction: A Convolutional Neural Network Approach", January2018.

[13] Jürgen Schmidhuber, "Deep learning in neural networks: An overview", January 2015.

[14] Breiman, L., Friedman, J., Stone, C.J., Olshen, R., "Classification and Regression Trees", 1st ed.; USA, 1984.

[15] M Kowsigan, P Balasubramanie- Volume No.22, Issue.5-An efficient performance evaluation model for the resource clusters in cloud environment using continuous time Markov chain and Poisson process, 2019.

[16] Thierry Bouwmans, Sajid Javed, Maryam Sultana, Soon Ki Jung, "Deep neural network concepts for background subtraction: A systematic review and comparative evaluation", September 2019.

[17] Kanika Mandal. "Predict IOT Sensor Failure", 2016.

[18] Bashir Mohammed, Irfan Awan, Hassan Ugail, Muhammad Younas, "Failure prediction using machine learning in a virtualised HPC system and application", June2019.

[19] M Kowsigan, A Christy Jebamalar, S Shobika, R

March-April 2020
ISSN: 0193-4120 Page No. 1461 - 1467

Roshini, A Saravanan-Volume No.14, Issue.2 Heart Disease Prediction by Analysing Various Parameters Using Fuzzy Logic, 2017.

[20] M Kowsigan, S Priyadharshini, N Sathish Kumar, C Vikramkumar-Volume No.118, Issue.18 -Security in Data and Dissemination of Distributed Data in Wireless Sensor Network, 2018.

[21] M Kowsigan, M Rubasri, R Sujithra, H Sumaiya Banu-Volume No.7, Issue.3 Data Security and Data Dissemination of Distributed Data in Wireless Sensor Networks, 2017.

[22] ACJ. Malar, M Kowsigan, N Krishnamoorthy, S Karthick, E Prabhu- Multi constraints applied energy efficient routing technique based on ant colony optimization used for disaster resilient location detection in mobile ad-hoc network.(2020).